2: Caroe S-4 (no date) 1953 RL1 Soviet Central Asia RM.

The central points in the earlier history of Turkistan are two. The first is the Mongol deluge which, under Chingiz Khan, swept over the land in the years 1219–24; the second the thunderbolt launched by Timur nearly two centuries later (1379–1405). Timur, though of Mongol antecedents, used Chaghatai Turkish as his mother tongue and belonged to what Barthold [1] called *la tribu mongole turquisée* of the Barlas, established at the ancient twin city of Kesh [2] where the Kashka Darya issues from the mountains.

It will be found that the human material, and the forces applied to it, will fall into a more comprehensible order if viewed in periods divided by these two outstanding events in the history of Central Asia. But, before embarking on any analysis of these periods, there are a number of general distinctions and definitions that I must make, if there is to be any understanding of the people mixed together in the crucible which is Central Asia.

A reference to the map at the end of this book will show the reader the disposition of the peoples of Russian Central Asia, arranged as it stands today in accordance with the various languages, Turkic and Iranian, spoken by the agglomerations known loosely as Uzbeks, Kazaks, Karakalpaks, Kirghiz, Turkmens (all Turkish) and the Tajiks (Iranians). It will be seen that in more than one case the speakers of these languages overlap across the international frontiers into China, Afghanistan and Persia. There are also areas within the country in which Uzbek mingles with Tajik, Turkmen or Karakalpak, Kirghiz with Uzbek, and so on. Bilingualism in Uzbek and Tajik is common in areas of overlap, especially so in the towns from Khojent to Bukhara. Turkic Uzbek in the towns has lost its vowel harmony under Iranian influence, but has preserved it in the north of the Uzbek linguistic territory, where there was no such influence. On the other hand, the juxta-

[1] *Op. cit.* p. 168.

[2] The modern name of these cities is Shahr-i-Sabz and Kitab. They are separated from Samarkand by the Takhta Karacha pass over a spur of the Hissar mountains.

2: Caroe S-4 (no date) 1953 RL1 Soviet Central Asia RL 1

Persian as the official language for administration and the law.[1] Literary Chaghatai, written in the Persian *shikasta* form of Arabic letters, has another and important liability, if regarded as a mode of transcription of a Turkic language. This is the relative weakness of the Arabic alphabet in vowel signs, and its consequent inability to express the very varied vowels of Turkic languages, and particularly the vowel qualities and harmonies. It is true that the form of Uzbek itself, as spoken at least in Tashkent, Bukhara and Samarkand, has become Persianized to the extent of losing its vowel quality and most of its vowel harmony. But, even so, Chaghatai Arabic expresses the vowel sounds inadequately, the more so because in the Farghana and Yese (Turkistan) forms of Uzbek the harmonies are still to be found. *A fortiori* Chaghatai, based (as Babur tells us) on the purest Uzbek, spoken then as today in the Andijan district of Farghana, was not a medium easily adaptable to the languages spoken in Kirghizia or in the countries now known as the Kazak and Turkmen steppe. There is an elaborate system of comparative morphology to be acquired before it becomes possible for a student, even on paper, to relate the various Eastern Turkic languages, *e.g.* Uzbek, Kazak and Turkmen, one to the other. In speech the difference is great enough to cause difficulty of understanding between even native speakers. The path of the reformer seeking to establish one language as current coin for all Turkistan is therefore not so easy as has often been suggested.[2] As a long-term project, if circumstances were favourable, it should not be impossible. But, as we know, the Soviets are doing all they can to promote linguistic and other distinctions between the five republics, and it follows that any move towards assimilation runs counter to tendencies within modern Turkistan itself. None the less, as a social and political

[1] Exactly the same process can be seen at work today in Afghanistan. The Barakzai Sardars who rule the country, themselves of Afghan blood, often can not even talk Pashtu, the Afghan tongue. Persian is the official language.

[2] A group of émigrés in Germany, mainly Uzbek, issue a journal, *Millij Türkistan*, in a language which they call *ortatili*, or the middle tongue. It is really Uzbek, with a few extraneous words, written in an alphabet adapted from the Turkish Roman. It is an interesting experiment.

2: Caroe S-4 (no date) 1953 RL1 Soviet Central Asia KL1

APPENDIX

TRANSLITERATION

A SYSTEM of transliteration of names should be consistent and broadly scholarly without being pedantic. It is particularly hard to observe this mean in the presentation of a Turkic nomenclature for English readers, and harder still when the theme is not Turkey but the Eastern Turks under Russian rule. The difficulty is still further increased by the fact that the Soviets have forced the Turkic and Iranian languages of Russian Central Asia into two distinct transcriptions, first a set of Roman letters, and then a set of Cyrillic letters. There have also been variants in the Roman and in the Cyrillic alphabets used to transcribe the various languages, *e.g.* a slightly different alphabet has been applied to Uzbek and Kazak. A final complication is introduced by my use of Togan's book in modern Western Turkish as one of the media through which these Trans-Caspian names are here presented for study. (It is consoling to observe that even Togan does not use an alphabet always consistent with that used in Turkey today, *e.g.* he introduces a '*q*' for Arabic ق, and a '*kh*' for Arabic خ, both unknown to the modern Turkish alphabet.)

2. Turkic nomenclature has often, by no means always, a Muslim background, but even then the value of the Arabic or Persian words is altered both in pronunciation and spelling, *e.g.* in the simplest case Mehmet for Muhammad. There are also the modifications (umlaut) on vowel sounds, and the vowel harmonies. In modern Turkish there is no '*kh*' (Russian *x*) for Arabic خ, its place being taken by a simple '*h*', but خ does exist in Eastern Turkic tongues. Finally, superimposed on the Persian and Arabic, one finds many names of pure Turkic origin.

2: Caroe S-4 (no date) 1953 RL1 Soviet Central Asia RL1

• 174 3. For these reasons the accepted Western system of transliteration adopted by Arabists does not fit any Turkic tongue. For instance, 'Abdu'l Raḥman or Ḥafidh, apart from their pedantic look, do not represent the words as they are apprehended by a Turk's eye or ear, west or east of the Caspian. There is rather more to be said for the simpler Hunterian system as used for Muhammadan nomenclature in India, but that too misses the distinctive elaboration of vowel sounds to be found in Turkic languages. It has, however, the advantage of familiarity to an English eye, and some adaptation of it may be a possible solution.

4. Next to be considered is a system taken through the Russian. This has the advantage that it represents the written word within modern Turkistan, and is therefore familiar to the educated in the five Central Asian republics. It is, however, anathema to every Eastern Turk outside the USSR and, opposed as it deliberately is to the memories and traditions of a Muslim past, is probably to the distaste of many of the indigenous populations within the country. It is, moreover, clumsy, *e.g.* the use of *dẓh* for Arabic ج (jim) and the Russian patronymic *-ov* so strange to a Muslim ear. Why write *Tadẓhik* in English for *Tajik*, or *Shayakhmetov* for the son of *Shah Ahmed*?

5. There remains the modern Western Turkish Roman alphabet. In many ways this might be suitable, but it involves the use of the following:

c for English *j* (French *dj*, Russian *дж*)
ç for English *ch* (French *tch*)
ş for English *sh* (French *ch*)
h for Arabic and Persian خ (German or Scottish *ch*, Russian *x*).

N.B.—In English systems this is usually given as *kh*. In modern Western Turkish it has lost its original sound and is pronounced as an aspirated *h* only. But the original sound remains east of the Caspian. In any case, words such as

2: Caroe S-4 (no date) 1953 RL1 Soviet Central Asia

Khiva, Bukhara, Khan look absurd in the dress *Hiva, Buhara, Han.*

ğ for Arabic and Persian غ (not unlike French *r*)

N.B.—In English systems usually given as *gh*.

k for Arabic ك (not unlike *ck* in cock)

N.B.—In English systems usually given as *q*.

ō as in German

ū as in German

ı (undotted) for Russian ы, something like English *u* in *hut* or *but*, or *y* in the adverbial termination *-ly*.

The use of all these in an English text would make it tiresome and unfamiliar to the general reader, and might introduce difficulty for the printer. All should be ruled out except *ō* and *ū* and possibly *k*.

6. It is easy to proceed so far negatively. As a positive contribution I suggest the following, which rejects the Russian and the scholarly Arabic, and represents a compromise between the modern Turkish and the Hunterian.

(*a*) *Vowels.* Adopt Western Turkish vowel signs except *ı* (undotted). (I have considered possibility of using *y* for *ı*, but this would be unsightly and unfamiliar, *e.g. Osmanly* for *Osmanlı*. On the whole it is probably better that *ı* (undotted) should appear without distinction from *i*.)

N.B.—I should modify this general principle in the case of names whose spelling has become standardized according to Hunterian values, *e.g.* write Samarkand, Farghana, Timur, Turk, not Semerkend, Ferghana, Temūr, Tūrk.

(*b*) *Consonants.* Adopt Western Turkish consonants with following important exceptions:

(i) Turkish *c* to become *j*, *e.g.* Tajik, not Tacik, Tadzhik or Tadjik;

(ii) Turkish *ç* to become *ch*, *e.g.* Chaghatai, not Çaghatai or Tchaghatai;

2: Caroe S-4 (no date) 1953 RL1 Soviet Central Asia RL1

● 194 (iii) Turkish ş to become *sh*, *e.g.* Shah, Sherif, not Şah, Şerif, Chah, Cherif;

(iv) Turkish *h*, when used for Persian or Arabic خ, to become *kh* (not *x*), *e.g.* Khan, Khiva, Sheikh, not Han, Hiva, Şeih;

(v) Turkish ğ to become *gh*, *e.g.* oghlu, not oğlu; Chaghatai, not Chagatai.

N.B.—I have considered what to do with Turkish *k* when used for ق. Scientific systems use *q*, but this would involve writing *e.g.* Qazaq, Qirghiz and Samarqand, all somewhat pedantic. Kazak and Kirghiz are probably better (Kazakh often used, following the Russian, is wrong in any case, the Arabic form being قزاق not قزاخ).

7. It is suggested that the text need not be complicated by using commas for *ain* and *hamza*. They are apt to give an air of dust, and are not very important for Turkic languages, for which is needed a system generally phonetic in intention and effect, yet conforming broadly with an educated approach to orientalia.

104 8. One name calls for comment. It may be said that Jengis, Jenghis, and Cheng(h)iz Khan are recognizable, but Chingiz is unfamiliar. On the system suggested Chingiz is a correct transliteration of the spelling adopted by Babur and others in Chaghatai. A correct transliteration from the original Mongolian would be Chingis, since there is no ڭ in the old Mongol writing.